

MANAGEMENT OF VIRTUAL TAPE VOLUMES USING DATA PAGE ATOMIC UNIT

Publication number: JP2003513352T

Publication date: 2003-04-08

Inventor:

Applicant:

Classification:

- international: G06F3/06; G06F12/00; G06F12/08; G06F3/06;
G06F12/00; G06F12/08; (IPC1-7): G06F12/00;
G06F3/06

- European: G06F3/06M; G06F12/08

Application number: JP20010533502T 20001024

Priority number(s): US19990426516 19991026; WO2000US29324
20001024

Also published as:

WO0131431 (A3)
WO0131431 (A2)
EP1226491 (A3)
EP1226491 (A2)
US6816957 (B1)

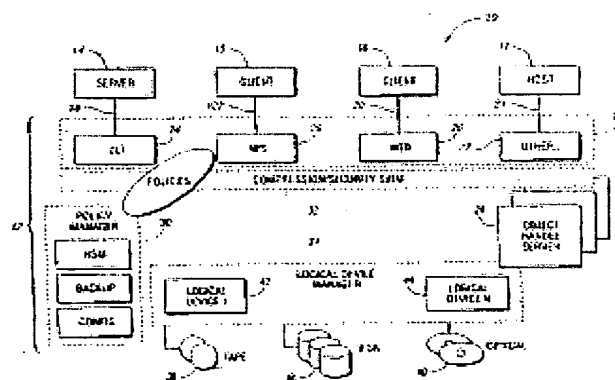
more >>

Report a data error he

Abstract not available for JP2003513352T

Abstract of corresponding document: **WO0131431**

An intelligent data storage manager is operational in a virtual data storage system that uses a storage element to store data for a host processor. The storage manager includes a host interface and a controller. The host interface receives a virtual volume having virtual blocks from a host processor. The controller sequentially segments the virtual volume into data pages. Each data page includes virtual blocks of the virtual volume. The controller generates a meta data page associated with the virtual volume to map the data pages to the virtual blocks of the virtual volume. The controller uses the meta data page to selectively store the data pages onto the storage element. The storage element has a partitioning configuration and the amount of the virtual blocks included by each data page is a function of the partitioning configuration of the storage element. The meta data page associated with the virtual volume includes meta data page records. Each meta data page record includes meta data records and each of the meta data records corresponds to at least one virtual block. The storage element includes high and low performance storage elements. The controller uses the meta data page to recall selected data pages from the low performance storage element onto the high performance storage element. The controller uses the meta data page to migrate selected data pages from the high performance storage element onto the low performance storage element.



Data supplied from the esp@cenet database - Worldwide

拒絶理由通知書

特許出願の番号 特願 2003-337239
起案日 平成19年12月11日
特許庁審査官 ▲吉▼田 美彦 9384 5100
特許出願人代理人 特許業務法人ウィルフォート国際特許事務所（
外 3名） 様
適用条文 第29条第2項、第36条

この出願は、次の理由によって拒絶をすべきものです。これについて意見がありましたら、この通知書の発送の日から60日以内に意見書を提出してください。

理 由

(1) この出願は、特許請求の範囲の記載が下記の点で、特許法第36条第6項第2号に規定する要件を満たしていない。

記

*請求項1、2、6、7

*備考

「前記第1の記憶制御装置は、前記上位装置によりアクセスされる少なくとも1つ以上の論理ユニットと、この論理ユニットと少なくとも1つ以上の記憶デバイスとを接続するようにして設けられる少なくとも1つ以上の中間記憶階層とを備えており、」の記載から、上位装置が認識する単位が論理ユニットであり、この論理ユニットと物理的な記憶デバイスを接続（マッピング？）するのが中間記憶階層であり、第1の記憶制御装置は該中間記憶階層を備えていることが読み取れる。

その上で、「前記中間階層のうち少なくとも1つは、前記第2の記憶制御装置が有する記憶デバイスに接続されている」の記載は、そもそも「前記中間階層」の前記には、中間記憶階層のみが記載され、中間階層なるものは記載無く、対応が不明確な上、記憶デバイスは、論理ユニットと接続されるものであるのに、中間階層に接続されとは、その関係が不明確である。

よって、発明を特定する構成要素間の関係が不明瞭で、発明が不明確。

請求項2、6、7の記載も同様に不明確。

*請求項 3

*備考

「前記中間記憶階層は、前記記憶デバイス上に設定される少なくとも1つ以上の第1の記憶階層と、この第1の記憶階層上に設定される少なくとも1つ以上の第2の記憶階層とを備えて構成されており、前記第2の記憶制御装置が有する記憶デバイスは、前記第1の記憶階層にマッピングされている」の記載では、第1の記憶階層上に設定されるとは如何なる技術的事項か不明な上、第2の記憶階層は何ら他の構成要素と関係が無く、請求項の記載自体、技術的に何を特定するのか不明瞭。

請求項の記載は、上位装置がアクセス対象とする論理ユニットに対して、如何に第1の記憶制御装置が、第1の記憶制御装置と第2の記憶制御装置が有する物理的記憶デバイスを、論理的に隠蔽してマッピングするか、発明を特定する構成要素間の関係が不明瞭で発明が不明確。

よって、請求項1～7に係る発明は明確でない。

(2) この出願の下記の請求項に係る発明は、その出願前に日本国内又は外国において、頒布された下記の刊行物に記載された発明又は電気通信回線を通じて公衆に利用可能となった発明に基いて、その出願前にその発明の属する技術の分野における通常の知識を有する者が容易に発明をすることができたものであるから、特許法第29条第2項の規定により特許を受けることができない。

記 (引用文献等については引用文献等一覧参照)

*請求項 1～7

*引用文献 1～2

*備考

請求項の記載は不明瞭な点があるものの、明細書を参酌し、上位装置が認識する論理ユニットに対して、中間的な論理装置をマッピングしておき、該中間的な論理装置に対する物理装置を第1の記憶制御装置が有する記憶デバイスか否かを問わずシームレスに接続可能なマッピング手法程度の事項として、判断する。

論理デバイスの設定として、階層記憶構成をもちいて、上位装置に対して下位の装置を仮想化することは例えば、引用文献1や引用文献2に記載されるように、上位装置の設定変更を無くす等のために、通常行われている設計事項であり、本願請求項1～7に係る発明は、引用文献1、2に対して、進歩性を有さない。

引用文献等一覧

1. 特表2003-513352号公報
 2. 特開2002-215334号公報
-

先行技術文献調査結果の記録

・調査した分野 I P C G 0 6 F 3 / 0 6 - 3 / 0 8

この先行技術文献調査結果の記録は拒絶理由を構成するものではありません。

この拒絶理由通知の内容に関するお問い合わせ、または面接のご希望がございましたら下記までご連絡下さい。

特許審査第4部 インターフェイス 吉田 美彦
TEL. 03 (3581) 1101 内線3565
FAX. 03 (3580) 6907

(11)特許出願公表番号
特表2003-513352
(P2003-513352A)

(43)公表日 平成15年4月8日(2003.4.8)

(51) Int.Cl. ⁷	識別記号	F I	デマコト* (参考)
G 0 6 F 12/00	5 1 4	G 0 6 F 12/00	5 1 4 E 5 B 0 6 5
	5 0 1		5 0 1 B 5 B 0 8 2
3/06	3 0 2	3/06	3 0 2 J
	3 0 3		3 0 3 Z

審查請求 有 予備審查請求 有 (全 26 頁)

(21) 出願番号	特願2001－533502(P2001－533502)
(86) (22) 出願日	平成12年10月24日(2000. 10. 24)
(85) 翻訳文提出日	平成14年2月15日(2002. 2. 15)
(86) 国際出願番号	PCT/US00/29324
(87) 国際公開番号	WO01/031431
(87) 国際公開日	平成13年5月3日(2001. 5. 3)
(31) 優先権主張番号	09/426, 516
(32) 優先日	平成11年10月26日(1999. 10. 26)
(33) 優先権主張国	米国(US)

(71)出願人 ストレイジ・テクノロジー・コーポレイション
アメリカ合衆国 80028-4309 コロラド州 ルイヴィル エムエス-4309 ワン
ストレイジテック ドライブ

(72)発明者 ステーパーン ハラディ
アメリカ合衆国 80027 コロラド州 ス
ベリオル モナーク ウェイ 1049

(72)発明者 デービッド エー メイジャー
アメリカ合衆国 80020 コロラド州 ブ
ルームフィールド チェリーウッド スト
リート 12163

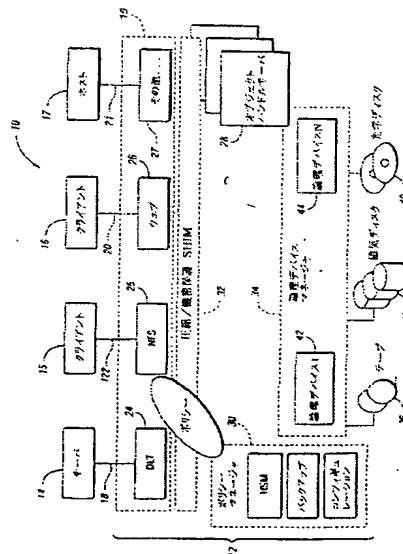
(74)代理人 弁理士 西山 善章

最終頁に続く

(54) 【発明の名称】 データページ・アトミックユニットを使用する仮想テープボリューム管理装置

(57) 【要約】

ホストプロセッサのためにデータを格納する記憶素子を使用する仮想データ記憶システムにおいて動作するインテリジェントデータ格納管理装置を提供する。本発明の格納管理装置はホストインターフェースとコントローラを含む。ホストインターフェースはホストプロセッサから仮想ブロックをもった仮想ボリュームを受信する。コントローラはその仮想ボリュームをシーケンシャルセグメントに分割してデータページにする。各データページはその仮想ボリュームの仮想ブロックを含む。コントローラは、データページをこの仮想ボリュームの仮想ブロックにマップするためのメタデータページを、仮想ボリュームに関連づけて発生する。コントローラはこのメタデータページを使用して、選択的にこれらデータページを記憶素子に格納する。記憶素子はパーティションコンフィギュレーションをもっており、各データページに含まれる仮想ブロックの量は記憶素子のパーティションコンフィギュレーションの関数である。仮想ボリュームに関連づけられたメタデータページはメタデータページレコードを含む。各メタデータページレコードはメタデ



【特許請求の範囲】

【請求項1】 データ記憶システム内で作動するインテリジェントデータ格納管理装置であって、該データ記憶システムが少なくとも一つのデータ記憶素子を使用して、該データ記憶システムに接続された少なくとも一つのホストプロセッサのために前記データ記憶素子にデータを記憶するようにされているインテリジェントデータ格納管理装置において、

該インテリジェントデータ格納管理装置は、

複数のブロックを有するボリュームをホストプロセッサから受信するホストインターフェースと、

前記ボリュームをシーケンシャルセグメントにセグメント化してデータページに分割するコントローラと、を含み、

前記データページの夫々が、前記ボリュームの少なくとも一つのブロックを含み、前記データページを前記仮想テープボリュームのそれぞれの仮想テープブロックにマップするためのメタデータページを、前記コントローラが前記仮想テープボリュームに関連づけて発生する、ことを特徴とするインテリジェントデータ格納管理装置。

【請求項2】 前記コントローラは、前記メタデータページを使用して選択的に前記データページを少なくとも一つのデータ記憶素子に格納することを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項3】 各前記少なくとも一つのデータ記憶素子がパーティションコンフィギュレーションを有し、

各データページに含まれる前記少なくとも一つのブロックの量が各前記少なくとも一つのデータ記憶素子の前記パーティションコンフィギュレーションの関数である

ことを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項4】 前記データページの夫々が、同一量のデータを含むことを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項5】 前記データページの夫々は、少なくとも200個のブロックを含むことを特徴とする請求項1に記載のインテリジェントデータ格納管理装置

【請求項6】 前記ボリュームがテープボリュームであり、前記ブロックがテープブロックであることを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項7】 前記ボリュームに関連づけられた前記メタデータページが複数のメタデータページレコードを含み、

各前記メタデータページレコードが複数のメタデータレコードを含み、

各前記メタデータレコードが少なくとも一つのブロックに対応することを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項8】 前記少なくとも一つのデータ記憶素子が少なくと一つの高パフォーマンスデータ記憶素子と少なくとも一つの低パフォーマンスデータ記憶素子とを含み、

前記コントローラが前記メタデータページを使用して、少なくとも一つの低パフォーマンスデータ記憶素子から、選択されたデータページを前記少なくとも一つの高パフォーマンスデータ記憶素子上に回収することを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項9】 前記少なくとも一つのデータ記憶素子が少なくと一つの高パフォーマンスデータ記憶素子と少なくとも一つの低パフォーマンス記憶素子とを含み、

前記コントローラが前記メタデータページを使用して前記少なくとも一つの高パフォーマンスデータ記憶素子から、選択されたデータページを前記少なくとも一つの低パフォーマンスデータ記憶素子へ移動することを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項10】 前記少なくとも一つのデータ記憶素子が、可変パフォーマンスをもつ複数のデータ記憶素子を含み、

前記コントローラが前記メタデータページを使用して前記データページを選択的に前記少なくとも一つのデータ記憶素子に格納することを特徴とする請求項1に記載のインテリジェントデータ格納管理装置。

【請求項11】 仮想データ記憶システム内で作動するインテリジェントデ

ータ格納管理装置であって、該仮想データ記憶システムが複数の高及び低パフォーマンスデータ記憶素子を使用して、該仮想データ記憶システムに接続された少なくとも一つのホストプロセッサのために前記データ記憶素子にデータを記憶するように構成されているインテリジェントデータ格納管理装置において、

該インテリジェントデータ格納管理装置が、

複数の仮想テープブロックを有する仮想テープボリュームをホストプロセッサから受信するホストインターフェースと、

前記ボリュームをシーケンシャルセグメントにセグメント化してデータページに分割するコントローラと

を含み、

各データページが前記仮想テープボリュームの少なくとも一つの仮想テープブロックを含み、

前記データページを前記仮想テープボリュームのそれぞれの仮想テープブロックにマップするためのメタデータページを、前記コントローラが前記仮想テープボリュームに関連づけて発生し、

前記コントローラは前記メタデータページを使用して前記データページを選択的に前記高及び低パフォーマンスデータ記憶素子に格納する、ことを特徴とするインテリジェントデータ格納管理装置。

【請求項12】 前記高及び低パフォーマンスデータ記憶素子が各々パーティションコンフィギュレーションを有し、

各データページに含まれる前記少なくとも一つの仮想テープブロックの量が前記高及び低パフォーマンスデータ記憶素子の少なくとも一つのパーティションコンフィギュレーションの関数である

ことを特徴とする請求項11に記載のインテリジェントデータ格納管理装置。

【請求項13】 前記データページの夫々が、同一量のデータを含むことを特徴とする請求項11に記載のインテリジェントデータ格納管理装置。

【請求項14】 前記仮想ボリュームに関連づけられた前記メタデータページが複数のメタデータページレコードを含み、

各前記メタデータページレコードが複数のメタデータレコードを含み、

各前記メタデータレコードが少なくとも一つの仮想テープブロックに対応することを特徴とする請求項11に記載のインテリジェントデータ格納管理装置。

【請求項15】 前記コントローラが前記メタデータページを使用して、前記少なくとも一つの低パフォーマンスデータ記憶素子から、選択したデータページを前記少なくとも一つの高パフォーマンスデータ記憶素子上に回収することを特徴とする請求項11に記載のインテリジェントデータ格納管理装置。

【請求項16】 前記コントローラが前記メタデータページを使用して、前記少なくとも一つの高パフォーマンスデータ記憶素子から、選択されたデータページを前記少なくとも一つの低パフォーマンスデータ記憶素子に移動することを特徴とする請求項11に記載のインテリジェントデータ格納管理装置。

【請求項17】 仮想データ記憶システムに接続された少なくとも一つのホストプロセッサのためにデータを記憶するためのデータ記憶素子を少なくとも一つ使用する仮想データ記憶システムにおいて動作するインテリジェントデータ格納管理装置を動作させる方法であって、

該方法が、

複数のブロックを有する仮想ボリュームをホストプロセッサから受信するステップと、

前記仮想ボリュームをシーケンシャルセグメントにセグメント化して、前記仮想ボリュームの少なくとも一つの仮想ブロックを含むデータページに分割するステップと、

前記データページを前記仮想ボリュームのそれぞれの仮想ブロックにマップするためのメタデータページを、前記仮想ボリュームに関連づけて発生するステップと、を含むことを特徴とするインテリジェントデータ格納管理装置を動作させる方法。

【請求項18】 前記データページを選択的に少なくとも一つのデータ記憶素子上に格納するために前記メタデータページを使用することを特徴とする請求項17に記載の方法。

【請求項19】 各前記少なくとも一つのデータ記憶素子がパーティションコンフィギュレーションを有し、

各データページに含まれる前記少なくとも一つの仮想ブロックの量が各前記少なくとも一つのデータ記憶素子のパーティションコンフィギュレーションの関数であることを特徴とする請求項 17 に記載の方法。

【請求項 20】 前記データページの夫々が、同一量のデータを含むことを特徴とする請求項 17 に記載の方法。

【請求項 21】 前記仮想ボリュームが仮想テープボリュームであり、前記仮想ブロックが仮想テープブロックであることを特徴とする請求項 17 に記載の方法。

【請求項 22】 前記仮想ボリュームに関連するメタデータページが複数のメタデータページレコードを含み、

各前記メタデータページレコードが複数のメタデータレコードを含み、

各前記メタデータレコードが少なくとも一つの仮想ブロックに対応することを特徴とする請求項 17 に記載の方法。

【請求項 23】 前記少なくとも一つのデータ記憶素子が少なくとも一つの高パフォーマンスデータ記憶素子と少なくとも一つの低パフォーマンスデータ記憶素子とを含み、

前記方法が、前記メタデータページを使用して前記少なくとも一つの低パフォーマンスデータ記憶素子から、選択されたデータページを前記少なくとも一つの高パフォーマンスデータ記憶素子に移動させるステップを含むことを特徴とする請求項 17 に記載の方法。

【請求項 24】 前記少なくとも一つのデータ記憶素子が少なくとも一つの高パフォーマンスデータ記憶素子と少なくとも一つの低パフォーマンスデータ記憶素子とを含み、

前記方法が、前記メタデータページを使用して前記少なくとも一つの高パフォーマンスデータ記憶素子から、選択されたデータページを前記少なくとも一つの低パフォーマンスデータ記憶素子に移動させるステップを含むことを特徴とする請求項 17 に記載の方法。

【請求項 25】 前記少なくとも一つのデータ記憶素子が可変パフォーマンスをもつ複数のデータ記憶素子を含み、

該方法が、前記データページを選択的に前記少なくとも一つのデータ記憶素子に格納すべく前記メタデータページを使用することを特徴とする請求項17に記載の方法。

【発明の詳細な説明】**【0001】****[技術分野]**

本発明は一般にデータ記憶システムに関し、特に仮想テープボリュームをデータページと呼ばれるアトミックユニットにセグメント化するように構成された仮想データ記憶システム用のインテリジェントデータ格納管理装置に関する。

【0002】**[背景技術]**

仮想データ記憶システムにおいては、格納ネットワークを介してクライアントアプリケーションからインテリジェントデータ格納管理装置に書き込まれる仮想テープボリュームの管理に顕著な問題がある。これらの問題のいくつかは、記憶装置のアクセスパフォーマンスとそのコストの点に問題がある。これらの問題に対処するためにインテリジェントデータ格納管理装置は、使用する記憶管理属性と記憶管理ポリシーに応じて異なる論理デバイスおよび物理デバイスへ仮想テープボリューム又はその一部を行き来させるための、ある形態の階層記憶管理 (hierarchical storage management: HSM) に頼ることができる。

【0003】

他の問題はシステム帯域などのシステムパフォーマンスに関するものである。仮想テープボリュームはしばしばディスクサブシステム上に設けられ、テープバッファとして使用される。入力チャンネルとこのディスクサブシステムとの間にはしばしば速度のミスマッチ (不一致) が存在する。実際、システム全体のパフォーマンスにとってディスクサブシステムへのシーケンシャル書き込みアクセスの最適化が決定的に重要である。同様にして、仮想テープボリュームがバックエンドすなわちテープサブシステムへ書き込まれるときも、速度マッチングが望まれる。

【0004】

大きなファイルの場合には他の問題が存在する。テープバックアップアプリケーションによってインテリジェントデータ格納管理装置に書き込まれる仮想テープボリュームは非常に大きいことがある。有限なサイズのディスクサブシステム

を用いる場合、記憶管理のユニットのサイズが仮想テープボリュームのサイズより小さくない限り、十分量に達しない限りの同時的仮想テープボリュームアクセスが可能である。

【0005】

これらの問題に対処するために必要とされているのは、仮想テープボリュームをより小さな記憶ユニットに分割して当該仮想データ記憶システム内のこれらのより小さな記憶ユニットを管理する、仮想データ記憶システム用インテリジェントデータ格納管理装置である。

【0006】

[発明が解決しようとする課題]

したがって本発明の課題は、仮想データ記憶システムにおいて記憶デバイスのパフォーマンスにとって最適なサイズにしたデータページと呼ばれるアトミックユニットに仮想テープボリュームを分割（セグメント化）するインテリジェントデータ格納管理装置を与えることである。

【0007】

本発明のもう一つの課題は、仮想テープボリュームを構成するデータページを、仮想データ記憶システム内の異なる記憶装置に選択的に格納するインテリジェントデータ格納管理装置を与えることである。

【0008】

本発明のさらに別の課題は、仮想テープボリュームをデータページに分割し、選択的にそれらデータページを異なる記憶デバイスに格納し、当該仮想テープボリュームのブロックを関連のデータページにマップするためのメタデータページを生成する、インテリジェントデータ格納管理装置を与えることである。

【0009】

上記の課題及びその他の課題を達成するため、本発明は、少なくとも一つのデータ記憶素子を使用しその上に仮想データ記憶システムに結合された少なくとも一つのホストプロセッサのためにデータを格納する、インテリジェントデータ格納管理装置を与える。本インテリジェントデータ格納管理装置はホストインターフェースとコントローラとを含む。ホストインターフェースは複数の仮想ブロッ

クを有する仮想ボリュームをホストプロセッサから受信する。コントローラは仮想ボリュームをシーケンシャルセグメントにセグメント化しデータページに分割する。各データページは仮想ボリュームの仮想ブロックを少なくとも一つ含む。コントローラは、データページをそれぞれの仮想ボリュームの仮想ブロックにマップするためのメタデータページを、仮想ボリュームに関連づけて発生する。

【0010】

コントローラはこのメタデータページを使用して選択的に、データページを少なくとも一つのデータ記憶素子に格納することが好ましい。前記少なくとも一つのデータ記憶素子は各々にパーティションコンフィギュレーションをもっており、各データページに含まれる前記少なくとも一つの仮想ブロックの量は、各前記少なくとも一つのデータ記憶素子のパーティションコンフィギュレーションの関数である。各データページは同一量の仮想ブロックを含むことが好ましい。

【0011】

仮想ボリュームに関連づけられたメタデータページは複数のメタデータページレコードを含む。各メタデータページレコードは複数のメタデータレコードを含み、各メタデータレコードは少なくとも一つの仮想ブロックに対応する。

【0012】

好ましい実施形態では、上記少なくとも一つのデータ記憶素子は少なくとも一つの高パフォーマンスデータ記憶素子と少なくとも一つの低パフォーマンスデータ記憶素子とを含む。コントローラはこのメタデータページを使用して、選択したデータページを少なくとも一つの低パフォーマンスデータ記憶素子から上記少なくとも一つの高パフォーマンスデータ記憶素子上に回収する。コントローラはまた、メタデータページを使用して、少なくとも一つの高パフォーマンスデータ記憶素子から少なくとも一つの低パフォーマンスデータ記憶素子に、選択されたデータページを移動させる。

【0013】

これに代わる実施形態では、少なくとも一つのデータ記憶素子が可変パフォーマンスをもつ複数のデータ記憶素子を含む。コントローラはメタデータページを使用してデータページを選択的に上記少なくとも一つのデータ記憶素子に格納す

る。

【0014】

さらに本発明は、上記課題及び別の課題に沿って上記インテリジェントデータ格納管理装置を実行する方法を与える。

【0015】

本発明の上記課題、その他の課題、特徴、利点は添付の図面を参照しつつ本発明の最良の実施形態に関する以下の詳細な説明から直ちに明らかであろう。

【0016】

[発明の好ましい実施の形態]

一般に本発明は、小規模コンピュータシステムインターフェース（SCSI）でデータ交換ができる任意のホストから任意のデータブロックを受信し及び返信することと、これらホストにさらなるソフトウェアを要求せずに知能的な形で且つ当該データ所有者又はホストには目に見えない形で、これらデータブロックを管理することとができるようにすることを意図している。本発明は、ファイルシステムのプレゼンテーション（呈示）、仮想ディスクのプレゼンテーション等を含む多重的且つ同時的なプレゼンテーションを提供する。

【0017】

図1は本発明のインテリジェントデータ格納管理装置12が実現される仮想データ記憶システム10のアーキテクチャ全体を示すブロック図である。仮想データ記憶システム10は、それぞれのデータチャンネル18-21によって複数のホストプロセッサ14-17に結合されている。データチャンネル18-21はホストインターフェース19で終端する。ホストインターフェース19はネームサーバ24-27のレイヤーを与え、ホストプロセッサ14-17に対する定義済みの実在する物理デバイスインターフェースの仮想体を呈示する。ホストプロセッサ14-17に関する限り、ネームサーバ24-27は実物理デバイスを実行する。ネームサーバ24-27はホストプロセッサ14-17から受信したユーザーデータをユーザーデータオブジェクトに変換する。ユーザーデータオブジェクトは基準フォーマットに変換することができ、あるいはバイナリフォーマットのままに残すことができる。

【0018】

オブジェクトハンドラサーバ28はオブジェクトハンドラを論理デバイスアドレスにマップするもので、一つのデータオブジェクトを多重に呈示することを許容する。オブジェクトハンドラサーバ28はユーザーオブジェクトを格納のため、データ空間にマップする。このマッピングは、ポリシー管理装置30にプログラムされたポリシーにより決定され、機密保護レイヤー32に従わされる。当該オブジェクト空間に対する持続的格納は論理デバイス管理装置34によって決定されるが、論理デバイス管理装置34はユーザーデータオブジェクトを格納するためのポリシーに基づいて一論理デバイスを割当てるか又は生成する。

【0019】

論理データ記憶装置は一つの複合デバイスであって、実の物理的データ記憶デバイスで構成される。そのようなデバイスは例えば、磁気テープサブシステム36、磁気ディスクサブシステム38、光学ディスクサブシステム40、RAID5ディスクアレイを含んだ「論理デバイス1」42のような論理デバイス、及びネットワーク接続を介して行われる論理デバイスへのアクセスのような別の論理デバイスへのアクセスをするソフトウェアを含んだ「論理デバイスN」44、または上記デバイスの組み合わせである。論理デバイスの定義（名前）は、持続的格納を連想させる実際のデバイスの性質を抽象している。

【0020】

ホストインターフェース19はその作用時、ホストプロセッサ14-17の一つから仮想テープボリュームを受信する。仮想テープボリュームは複数の仮想テープブロックを含んでいる。各仮想テープブロックのサイズは可変である。一例を挙げると、各仮想テープブロックのサイズは32キロバイト程度である。

【0021】

各ネームサーバ24-27は仮想テープボリュームをシーケンシャルセグメントにセグメント化して（シーケンシャルセグメント化という）データページと呼ばれるアトミックユニットに分割する。各データページは少なくとも一つの仮想テープブロックを含む。ネームサーバ24-27はデータ記憶デバイスへのデータページの格納を制御することによってデータ記憶デバイスへの仮想テープボリ

ュームの格納を制御する。ネームサーバ24-27はホストプロセッサ14-17とデータ記憶デバイスとの間の入力／出力を強化するためにデータ記憶デバイスへのデータページの格納を制御する。

【0022】

各データページは一定のサイズをもっている。一例を挙げると、各データページはサイズが64メガバイト程度であり、200個の32キロバイトサイズの仮想テープブロックを含んでいる。各データページのサイズは、磁気テープサブシステム36及び磁気ディスクサブシステム38等のデータ記憶素子のコンフィギュレーションの関数である。例えば、インテリジェントデータ格納管理装置12が第三の記憶装置として磁気テープサブシステムを使用するのであれば、この場合これらのテープサブシステムは各パーティションが64メガバイトのサイズになるパーティション様式をうまく利用するので、このときはデータページも64メガバイトのサイズをもつようにすることにより、テープの空間利用に対して特定の利点を実現される。特に、「挿入編集」がイネーブル化されている場合は、データページをテープ上で有効期限順あるいは空間利用オペレーション順に並べる必要はない。

【0023】

インテリジェントデータ格納管理装置12に仮想テープボリュームが書き込まれるので、仮想テープボリュームを構成するデータストリームはネームサーバ24-27によってシーケンシャルセグメント化され、データページに分割される。各データページは複数の仮想テープブロックを含み、各仮想テープボリュームは複数のデータページを含んでいる。ネームサーバ24-27は、到来する仮想テープブロックをデータページにマップし、出て行く仮想テープブロックをデータページからマップする。

【0024】

データページはネームサーバ24-27と磁気テープサブシステム36及び磁気ディスクサブシステム38との間の転送ユニットである。磁気テープサブシステム36は、それが比較的到低コストであり、且つ磁気ディスクサブシステム38と較べて比較的に遅い点で、低パフォーマンスデバイスであると考えられる。

磁気ディスクサブシステム38は、それが比較的に高コストであり、且つ磁気テープサブシステム36と較べて比較的に高速である点から、高パフォーマンスデバイスであると考えられる。

【0025】

仮想テープボリュームをデータページに分割した後、それぞれのネームサーバ24-27はオブジェクトハンドルサーバ28とポリシー管理装置30と協力してホストインターフェース19から磁気ディスクサブシステム38上へ移されるデータページの流れを制御する。次いでそれぞれのネームサーバ24-27は再び、オブジェクトハンドルサーバ28およびポリシー管理装置30と協力して、特にポリシー管理装置のHSMコンポーネントと協力して、磁気ディスクサブシステム38から磁気テープサブシステム36への選択されたデータページの移動を制御する。データページは、最非近時使用 (least recently used) アルゴリズム、サイズで重みを付けた (size-weighted) アルゴリズムなどに基づいて、磁気ディスクサブシステム38から磁気テープサブシステム36へ移動するものが選択される。

【0026】

他の仮想テープボリューム、あるいはポリシーインターフェースを介して設定されるようなクライアント属性が、移動に影響を与えることがあり得る。例えば、ファイルシステム プレゼンテーションでは、ポリシーはあるデータページが永久的に高パフォーマンス記憶装置に常駐することを指示することができる。同様に、データページは、高いデータ信頼性の獲得、データ共有、および不意の故障からの回復力を発揮するために同時的に複数ロケーションに常駐することができる。仮想テープボリュームのデータページの集まりは磁気テープサブシステム36内の同一の物理的テープカートリッジあるいは多数ボリュームカートリッジへ移動する、という保証は全く無い。仮想テープボリュームあるいはクライアントレベルアフィニティ/非アフィニティポリシーを適用することは、移動と多数ボリュームカートリッジの空間利用アルゴリズムとに影響を与える。

【0027】

仮想テープボリュームから選択されたいくつかのデータページは磁気ディスク

サブシステム38上に置かれる。これらのデータページは「開かれて」いて、クライアントアプリケーションにより読み取り及び書き込みの予約がされている、と考えられる。仮想テープボリュームの他のデータページは、磁気ディスクサブシステム38からの移動後、磁気テープサブシステム36上に置かれる。磁気テープサブシステム36上に常駐するこれらのデータページは、クライアントアプリケーションがリクエストしたときに磁気ディスクサブシステム38上に回収して利用可能である。磁気ディスクサブシステム38が最適に使用されるように、すなわちクライアントアプリケーションがディスクサブシステムデータ回収パフォーマンスを享受できるように、ネームサーバ24-27はオブジェクトハンドラサーバ28とポリシー管理装置30の協力の下に適切な時点でインテリジェンスを使い、仮想テープボリュームのデータページのプレフェッチを行う。他方、仮想テープボリュームのデータページの大多数は、他の記憶階層の何処かに常駐する。

【0028】

ここで引き続き図1を参照すると共に図2乃至5を参照する。ネームサーバ24-27はデータページを仮想テープボリュームの仮想テープブロックにマップするためのメタデータページを発生する。図2はある仮想テープボリューム内のデータページの組織を表すツリー構造40の例を示す。オブジェクトハンドラサーバ28は本質的に、ツリー構造40によって示される階層ファイルシステムである。ツリー構造40は第一レベルでいろいろのクライアント登録簿42に分割されている。各クライアント登録簿42はいろいろの仮想テープボリューム44に分割される。各仮想テープボリューム44はデータページ46に分割されており、関連のメタデータページ表50を有する。

【0029】

図3は、メタデータページレコードコラム52とそれに対応するデータページコラム54とを有するメタデータページ表50を示す。仮想テープボリュームに対する各メタデータページは、メタデータヘッダ54と複数のデータレコード56とからなる。各データレコード56は、データページを構成する各仮想テープブロック集合に対応する。例えば図3に示すように、データレコード1は仮想テ

ープブロック集合1に対応する。データレコード56は仮想テープボリュームのデータページに対応する。例えば、データレコード1はデータページ1に対応し、データレコード2乃至nはデータページ2に対応する。

【0030】

データレコードn+1はファイルマークであり、対応するデータページをもたない。データレコード1は一定サイズをもつ仮想テープブロックを含む。データレコード2乃至nは可変サイズの仮想テープブロックの集団を含む。データレコード2乃至nは仮想テープブロックの集合2乃至nに対応する。

【0031】

一般にコントローラはメタデータページを使用して選択的にデータページを、インテリジェントデータ格納管理装置12と共働する記憶素子に格納する。高パフォーマンスをもつディスクサブシステム38と低パフォーマンスをもつテープサブシステム36とを有する記憶素子コンフィギュレーションでは、コントローラはメタデータページを使用して、選択されたデータページを磁気ディスクサブシステム38から磁気テープサブシステム36へ移動する。同様に、コントローラはメタデータページを使用して、仮想テープボリュームのうちの選択されたデータページを磁気テープサブシステム36から磁気ディスクサブシステム38上に回収する。

【0032】

図4はメタデータヘッダレコード60に含まれる情報を例示する。メタデータヘッダレコード60は一般に、クライアントアプリケーションに対するインテリジェントデータ格納管理装置12のプレゼンテーションに特有な制御情報を含む。メタデータヘッダレコード60もまた書き込み保護ステータスのような補助的情報を含む。図4に示すように、仮想テープボリュームに対するメタデータヘッダレコード60はバージョン番号情報62、作成日情報64、作成時刻情報66、デバイスタイプ情報68、媒体タイプ情報70、媒体密度情報72、およびファイル長情報74を含む。媒体タイプ情報70は仮想テープボリュームのデータページが書き込める媒体タイプと、データページの読み取り／書き込みアクセスのタイプとを記述する。ファイル長情報74はメタデータページの長さを記述す

る。

【0033】

図5はメタデータレコード80を例示する。メタデータレコード80は一般に、データレコードに特有な情報を含む。図5に示すように、メタデータレコード80はオフセット情報82、テープ残量情報84、連続ブロックカウンタ情報86、データブロックサイズ情報88、およびデータページ番号情報90を含む。連続ブロックカウンタ情報86およびデータブロックサイズ情報88はあるデータページ内のある特定の仮想テープブロックのロケーションを決定することを可能にする。

【0034】

上に述べたように、仮想テープブロックは可変サイズを有する。一定サイズの連続ブロックは、コントローラによって課されるデータページサイズの限界に達するまで、一メタデータレコード内で（従って連続ブロックカウンタ情報86及びデータブロックサイズ情報88内で）編成される。総てのブロックが一定サイズをもっているなら、各データページごとに一メタデータレコードが在ることになる。ブロックサイズが可変である最悪の場合、すなわち一定サイズの連続ブロックが全く無い場合、各ブロックごとに一メタデータレコードが存在する。上に指摘したように、いくつかのメタデータレコードは対応するデータページを全くもたない。とくにファイルマークを記録するメタデータレコードはそうである。

【0035】

特定アプリケーションのヘッダ情報とトレーラ（後書き）情報とを含む磁気テープサブシステム36にはある種のデータページを保持するオプションがある。このオプションは、低待ち時間型のテープサブシステム36から特定クライアントの情報を回収することを要せずに仮想テープボリュームマウント確認オペレーションが続行できることを保証する。

【0036】

上述したように、インテリジェントデータ格納管理装置12は、クライアントホストプロセッサ14-17とデータ記憶素子素子との間の入力／出力にデータ

ページを使用する。また、データページは制御フローにも使用できる。例えば、諸クライアントがシステムソースを求めて競合するような高負荷状態の下で、クライアントがインテリジェントデータ格納管理装置12を消費するのを統制するのにデータページの境界を使用することができる。すなわち、第一クライアントがデータページ分の入力／出力を完了するのを許され、その後、第一クライアントのプロセスは、第二クライアントがデータページ分の入力／出力を完了する間は保留される。これはマルチスレッドアルゴリズムに類似しており、いずれのクライアントアプリケーションも入力／出力を待っているうちにタイムアウト(時間切れ)のために入出力に失敗することを防止するのに採用できよう。

【0037】

高および低パフォーマンス記憶素子はインテリジェントデータ格納管理装置12と関連づけられていることが好ましい。それに代わる実施形態では、インテリジェントデータ格納管理装置12を同じパフォーマンス品質をもつ一つあるいは一組の記憶装置と関連づけてもよい。例えば、インテリジェントデータ格納管理装置12にテープデバイス形式のデータ記憶装置のみを関連づけることができる。この場合、仮想テープボリュームは(この場合、サイズにして256キロバイトで)データページ中に分散され、他のデータページと多重化されて単一の実テープボリュームに載せられる。この場合、データページの256キロバイトというサイズは主メモリバッファのパフォーマンスを増大するために選択されたものである。この場合、データページのサイズ256キロバイトは、主メモリバッファのパフォーマンスを大きくするように選択されたものである。これに代わる実施形態では、インテリジェントデータ格納管理装置12を、格納パフォーマンスが三以上のレベルすなわち階層をもつデータ格納コンフィギュレーションに関連づけることができる。

【0038】

仮想テープはインテリジェントデータ格納管理装置12の一つの表れに過ぎないから、これに代わる諸表示を考慮するときは、上述したテープブロックのデータページへの関連性の重要度は減じる。クライアントがネットワークファイルシステム又はウェブサービスを充足するためにインテリジェントデータ格納管理装

置12を使用しているときでも、データページは、一以上のファイルに関連づけられる「エクステンツ」(ブロックの一集合)に非常によく適合する。

【0039】

以上のように、本発明に基づいて前記の諸課題および利点を十分に満たす仮想データ記憶システム用のインテリジェントデータ格納管理装置が与えられることが明らかであろう。本発明は特定の実施形態について説明したが、当業者には多数の代替の実施形態、設計変更、変種があり得ることが明白であろう。従って、それら総ての実施形態が本発明の要旨および前記特許請求の範囲に含まれると理解されたい。

【図面の簡単な説明】

【図1】 本発明のインテリジェントデータ格納管理装置が用意された仮想データ記憶システムのアーキテクチャ全体を示すブロック図である。

【図2】 種々のクライアントアプリケーション用仮想テープボリュームのデータページの組織図である。

【図3】 メタデータページ表の例を示す図である。

【図4】 メタデータヘッダレコードの例を示す図である。

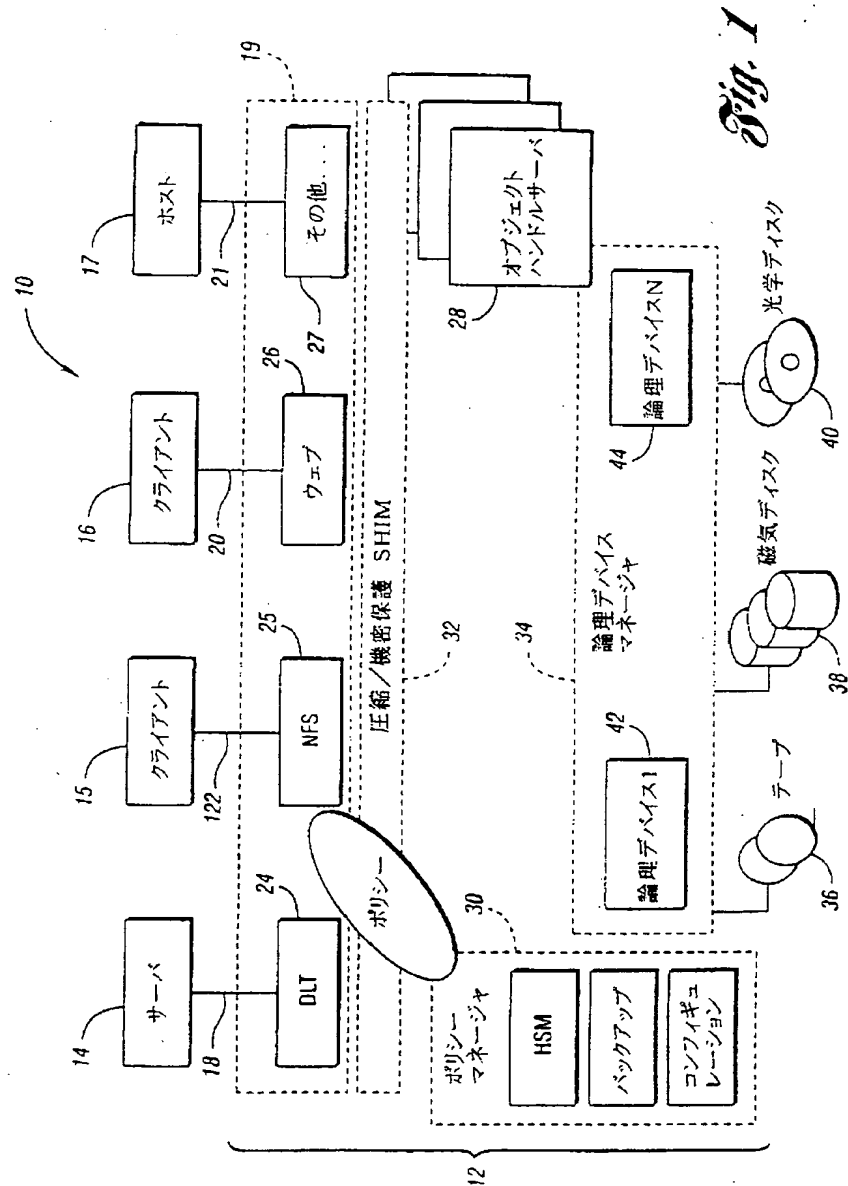
【図5】 メタデータレコードを例示する図である。

【符号の説明】

- | | |
|-------|-------------------|
| 10 | 仮想データ記憶システム |
| 12 | インテリジェントデータ格納管理装置 |
| 14～17 | ホストプロセッサ |
| 18～21 | データチャンネル |
| 24～27 | ネームサーバ |
| 28 | オブジェクトハンドルサーバ |
| 30 | ポリシー管理装置 |
| 32 | 機密保護レイヤー |
| 34 | 論理デバイス管理装置 |
| 36 | 磁気テープサブシステム |
| 38 | 磁気ディスクサブシステム |

- 4 0 光学ディスクサブシステム
- 4 2 論理デバイス 1
- 4 4 論理デバイス N
- H S M 階層記憶管理
- 4 6 データページ
- 5 0 メタデータページ表
- 5 2 メタデータページレコードコラム
- 5 4 データページコラム
- 5 4 メタデータヘッダ
- 5 6 データレコード
- 6 0 メタデータヘッダレコード
- 6 2 バージョン番号情報
- 6 4 作成日情報
- 6 6 作成時刻情報
- 6 8 デバイスタイプ情報
- 7 0 媒体タイプ情報
- 7 2 媒体密度情報
- 7 4 ファイル長情報
- 8 0 メタデータレコード
- 8 2 オフセット情報
- 8 4 テープ残量情報
- 8 6 連続ブロックカウンタ情報
- 8 6 データブロックサイズ情報
- 9 0 データページ番号情報

【図1】



【図2】

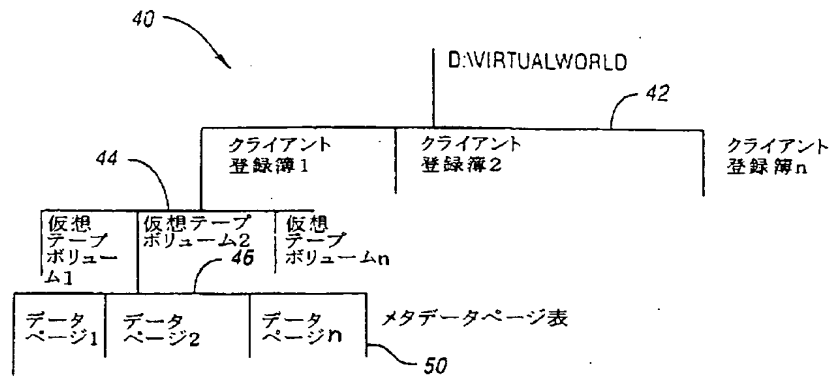


Fig. 2

【図3】

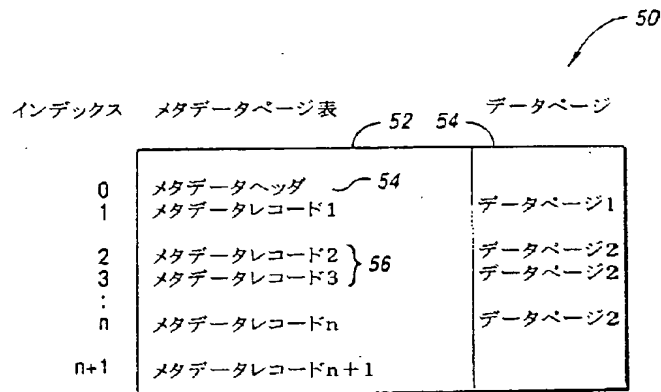


Fig. 3

【図 4】

情報内容	ビット	フォーマット	データタイプ	サイズ
62 ～ バージョン番号情報		"1.00"	Character	4 bytes
64 ～ 作成日情報		"mm/dd/yyyy"	String	Character 10 bytes
66 ～ 作成時刻情報		"hh:mm:ss"	String	Character 8 bytes
68 ～ デバイスタイプ情報	31	DLT7000		
	30 - 16	Reserved for future devices		
	15	SCSI-2		
	14 - 0	Reserved for future channel types		
70 ～ 媒体タイプ情報	6 - 0	DLT7000	Byte	1 byte
	7	85 Type IV Media		
		Cartridge Write Protection		
		0 Read/Write Access		
		1 Read Only Access		
72 ～ 媒体密度情報	7 - 0	DLT7000	Byte	1 byte
		18 35 GB without compression		
		82 20 GB without compression		
		83 40 GB with compression		
		84 35 GB without compression		
		85 70 GB with compression		
74 ～ ファイル長情報	31 - 0	Length of this Meta Data Page Table in bytes	Integer	4 bytes

Fig. 4

【図 5】

情報内容	ビット	フォーマット	データタイプ	サイズ
82 ～ オフセット情報	31 - 0	Absolute location of initial block in a data page	Integer	4 bytes
84 ～ テープ残量情報	63 - 0	Actual byte count of virtual tape remaining	Long Integer	8 bytes
86 ～ 連続ブロックカウンタ情報	31 - 0	Consecutive virtual tape blocks or filemarks	Integer	4 bytes
88 ～ データブロックサイズ情報	31 - 0	Virtual tape block size	Integer	4 bytes
90 ～ データページ番号情報	31	0 Data Compression On	Integer	4 bytes
		1 Data Compression OFF		
	30 - 0	Number 0 Filemark		
		1 - a Data Page Number		

Fig. 5

【国際調査報告】

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 G06F3/06 G06F12/08		International Application No. PCT/US 00/29324
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EP0-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4 467 421 A (WHITE BARRY B) 21 August 1984 (1984-08-21) column 9, line 33 - line 63 column 16, line 50 - column 18, line 8; figures 2-11	1-25
A	WO 99 03098 A (SUTMYN STORAGE CORP) 21 January 1999 (1999-01-21) page 3, line 19 - page 5, line 31; figures 1A, 1B	1, 11, 17
<input type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : 'A' document defining the general state of the art which is not considered to be of particular relevance 'E' earlier document but published on or after the international filing date 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim or other special reason (as specified) 'C' documents referring to an oral disclosure, use, exhibition or other means 'P' document published prior to the international filing date but later than the priority date claimed 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention 'X' document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone 'Y' document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art 'B' document member of the same patent family		
Date of the actual completion of the international search 16 May 2001		Date of mailing of the international search report 23/05/2001
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Moens, R

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 00/29324

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 4467421 A	21-08-1984	US RE36989 E	12-12-2000
		AU 6240780 A	30-04-1981
		BE 885727 A	15-04-1981
		CA 1153126 A	30-08-1983
		DE 3039285 A	30-04-1981
		FR 2468163 A	30-04-1981
		GB 2063532 A, B	03-06-1981
		IT 1132962 B	09-07-1986
		JP 1772143 C	14-07-1993
		JP 4048250 B	06-08-1992
		JP 56094452 A	30-07-1981
		MX 148922 A	07-07-1983
		NL 8004884 A	22-04-1981
		SE 8007292 A	19-04-1981
		CA 1174373 A	11-09-1984
		JP 4030057 B	20-05-1992
		JP 57195376 A	01-12-1982
WO 9903098 A	21-01-1999	AU 8390498 A	08-02-1999
		EP 1051704 A	15-11-2000

フロントページの続き

(81)指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW

(72)発明者 ロバート ビー ウッド

アメリカ合衆国 80503 コロラド州 ニ
ウォット ベッパーツリー ドライブ
6861

Fターム(参考) 5B065 BA07 CE04 CH18 ZA16
5B082 CA01 CA11 CA13 FA04

【要約の続き】

ータレコードを含み、各メタデータレコードは少なくとも一つの仮想ブロックに対応する。データ記憶素子は高及び低パフォーマンス記憶素子を含む。コントローラはこのメタデータページを使用して、選択したデータページを低パフォーマンスデータ記憶素子から高パフォーマンスデータ記憶素子上に回収する。コントローラはメタデータページを使用して、高パフォーマンスデータ記憶素子から低パフォーマンスデータ記憶素子に、選択されたデータページを移動させる。